ABSTRACT

Title: Measuring Fatigue-related Impairment in the Workplace

Authors: Ronald I Powell and Alex G Copping

Journal: Journal of Engineering, Design and Technology

Year: Published in 2016, Vol.14, No.3, pp507-525

Purpose - This research was founded on the premise that the safety record in the global construction industry needs improvement and that worker fatigue-impairment may be an underlying cause or major contributor to accidents. There have been few attempts to understand how pervasive and serious fatigue-impairment is in construction even though research has shown fatigue-impairment can be as big a concern as alcohol-impairment. When fatigue-impairment is acknowledged as existing, there is poor understanding of its severity and how it may affect many factors that contribute to accidents.

Design - This research built on actual measurements of fatigue-related impairment from workers on a large construction project which showed significant fatigue-related impairment. The research identified and tested possible techniques for real-time measurement solutions to assist with this safety-related issue.

Findings - Performance results from cognitive tests were compared with modeled mental effectiveness from actigraph-monitored sleep of 100 participants for a month each and showed significant correlation for all cognitive tests used.

Practical Implications - Derived from operational settings, this research confirmed the possible use of simple, quick and inexpensive cognitive tests as screening techniques for workplace impairment which could be used as part of a fatigue management plan.

Originality/value - This research confirms the need for and a solution for fatigue monitoring in the workplace.

Keywords: actigraph; cognitive tests; fatigue; impairment; sleep

INTRODUCTION

Construction is a high hazard occupation (National Institute for Occupational Safety and Health, 2007). Factors such as poor working conditions due to inclement weather, mobile equipment, travel to multiple and different work sites, changing and demanding schedules, long work hours, inadequate training and impaired workers contribute to this risky work environment. A risky environment has driven a safety focus for improvements but not all factors that may help the industry's safety and resulting productivity have been studied in detail such as fatigue-related impairment. An understanding of the degree fatigue-impairment may even be a workplace issue is underdeveloped. An important premise of this research was that this understanding is best addressed by measuring fatigue-impairment in operational settings.

An extensive study (Powell, 2009, Powell and Copping, 2010) was conducted on a group of construction workers associated with a large multi-faceted construction project in Canada. The study found the workers on average were getting inadequate sleep resulting in decrements in performance solely due to inadequate sleep. Actual sleep measurements with actigraphs of construction workers fed an impairment model and coupled with other relevant indicators pointed to a fatigue-impaired workforce with higher workplace accident risk. In construction work which exposes workers to many hazards and relies on alert workers to avoid accidents at all times, any impairment is concerning. Simple monitoring of workers is the primary means to identify all forms of impairment from which additional testing may be conducted if required. However, this is not the case with fatigue-related impairment which has no workplace tools to assist with its identification and control. Even if identified by monitoring, nothing is typically done about fatigue-impairment in the workplace as it carries little concern. This research addressed questions associated with whether tools could be used to identify and classify fatigue-impairment in real time to assist controlling it. Tools have been identified which purport to have this ability but they have not been used or tested in an operational setting with uncontrolled factors affecting individuals and their alertness. In particular, cognitive tests were identified as possible surrogate measurements for fatigue-related impairment. To assess the possibility of their use, output from the tests was compared to another proven standard which were estimations of mental effectiveness from an advanced actigraph-fed sleep-based model.

Sleep and Fatigue

Fatigue is a longer term condition than sleepiness and inadequate sleep versus an individual's need is considered to be a primary cause of fatigue-impairment (Beaulieu, 2005). Strauss (2003) who set early standards studying the effects of fatigue in real world aviation environments, defined fatigue as "a non-pathological state resulting in a decreased ability to maintain function or workload due to mental or physical stress" (ibid, p1). Strauss further believed that there are "…two major physiological phenomena that have been demonstrated to create fatigue: sleep loss and circadian rhythm disruption" (ibid, p1). The consequences of fatigue-related impairment may include a general diminishment in certain cognitive functions that affect tasks performed. It is most evident if someone falls asleep but most often the result is more subtle or undetectable such as slowed reaction time or loss of attentiveness. Workplace fatigue-related safety concerns are a focal point of this research but fatigue-related impairment does not have boundaries restricting it to a worksite.

Fatigue-impairment has safety consequences in our personal and work lives, including commuting.

A basic understanding of sleep is foundational to understanding fatigue and while we do not fully understand all aspects of sleep, we do know that regardless of how one has become fatigued, sleep is the only naturally occurring cure for it. Research continues on the fundamentals of sleep but relative to understanding sleep's contribution to workplace fatigue-impairment, there is ample understanding and an abundance of literature that illustrates the average amount of sleep needed by healthy adult populations is around 8 hours per 24 hours for full performance restoration (Levine *et al.*,1988, Wehr *et al.*,1993, Dinges *et al.*,1996, Rosekind *et al.*, 2000, Van Dongen *et al.*, 2003, Roth, 2006). Several studies suggest we get seven or less hours of sleep per 24-hour period resulting in a sleep deficit (Dinges *et al.,*1997). Ideal sleep varies for each individual and relative to meeting one's sleep requirements there is an important quality aspect. Quality of sleep is a measure of how closely an individual follows an uninterrupted natural sleep pattern. It has two components; one associated with the number of wake periods and sleep loss due to wake periods and a second associated with which stage of the sleep cycle the loss occurs in. It is normal to wake during the night and multiple times, but excessive waking during a night's sleep contributes no sleep value.

Sleep requirements are pre-programmed and governed according to the circadian cycle. Sleep follows our circadian rhythms and it is an important aspect of understanding the relationship between sleep, natural alertness levels and worker performance which has peaks and troughs over the course of a day as shown in Figure 1. It is a key factor especially for shift-workers.

<Insert Figure 1 here>

Sleepiness can be most problematic if it comes at an inopportune time relative to mental or physical demands. In a critical, demanding, or high-risk activity the outcome can be very consequential. Anyone for any reason not meeting their sleep requirements at night will exhibit a level of fatigue-related impairment. Fatigue-impairment rises and falls in the wake cycle. Fatigue gets imported into the workplace as a hazard and has been shown to have contributed to a large number of accidents and deaths. It can be both a driver and weakened defence for an accident at the site level. It can align with active failures (Reason, 2008) to cause an unsafe act and it can also weaken defences of another in harm's way. It must be noted that fatigue, like other impairing agents such as alcohol and drugs, are not narrow in their impact. They can influence multiple areas of cognition and weaken the defences to hazards. In the design of a construction project and worksite, factors such as staffing levels and overtime can lead to latent conditions of fatigue in workers that may lead to slowed reaction times and poor judgment. One worker's departure from safe work procedures may require another's vigilance and fast reaction time to avoid an accident, but fatigue weakens that defensive layer.

Measuring Sleep and Fatigue

Dawson and Reid's (Dawson and Reid, 1997) pioneering study which compared performance changes due to alcohol-impairment and sleep deprivation, allowed a pragmatic leap forward in quantifying fatigue-impairment. Several research projects in the past decade have refined this area resulting in correlations between lack of sleep and Blood Alcohol Concentration (BAC) levels (Lamond and Dawson, 1999, Falleti *et al.,* 2003, Fletcher *et al.,* 2003, Maruff *et al.,* 2005). Dawson and Reid's work is affiliated with two separate technologies used to assess fatigue-related performance. The first is cognitive testing which was used to replicate the findings of Dawson and Reid's correlation of time awake to BAC levels (Falleti *et al.,* 2003). While fatigue is not directly measurable there are detectable symptoms of fatigue such as degraded mental performance. A possible simple solution for detecting fatigue-related impairment is to test aspects of an individual's cognition. Such tests can be structured to assess different cognitive functions. Some of these tests may be more or less sensitive to fatigue. Common tests include attention tests, judgment tests, reaction time tests and memory tests. Memory tests can be designed to test short term, working and long term memories and have assisted in diagnosing Alzheimer's disease (Ellis *et al.,* 2009, Darby *et al.,* 2011). Applications for medical clinical use have been studied and abnormal conditions associated with attention may also lead to diagnosis of attention deficit disorder (Collie *et al.,* 2007).

The computerized cognitive tests chosen consisted of cognitive tasks known to be influenced by impairment and in particular, fatigue. The tasks use a game-like format to assess the cognitive domains of psychomotor processing, visual attention, learning and working memory. These domains have previously been shown sensitive to the effects of mild head injury and concussion, (Makdissi *et al.,* 2001, Collie, *et al.,* 2003, Moriarity *et al.,* 2004, Collie *et al.,* 2006, Maruff *et al.,* 2009, Straume-Naesheim *et al.,* 2009, Makdissi *et al.,* 2010) as well as fatigue and drug use, (Falleti *et al.,* 2003, Snyder, Bednar*, et al.,* 2005, Snyder, Werth*, et al.,* 2005, Collie *et al.,* 2007) psychiatric (Pietrzak, Olver*, et al.,* 2009, Pietrzak, Snyder*, et al.,* 2009) and neurodegenerative disease (Darby *et al.,* 2002, Maruff *et al.,* 2004, Ellis *et al.,* 2009, Darby *et al.,* 2011). Similar to other computerized approaches, each task requires a “YES” or “NO” response to stimuli, which in this case is upon display of a central playing card. Each time the test is taken, the visuomotor requirements remain identical, but equivalent alternate forms of each task are randomly generated. Reliability, stability, practice effects, validity and correlations with conventional neuropsychological tests have been reported previously, (Falleti *et al.,* 2006, Maruff *et al.,* 2009, Fredrickson *et al.,* 2010) as have other psychometric properties (Collie, Maruff, Mcstephen*, et al.,* 2003) which drives the interest in researching whether these simple, brief tests could be applied with success in a real world work environment to generate a declaration of fatigue status.

The tests sensitive to fatigue (Falleti *et al.,* 2003, Collie *et al.,* 2007) were selected to make up a battery to be completed at one sitting. The first was Detection (DET) which is a simple reaction-time based test requiring participants to monitor a card presented face down on the computer monitor and hit the ‘K’ key on the computer key board to indicate that “YES” it had turned over. The second was Identification (IDN), which is a choice reaction-time test requiring participants to monitor a card presented face down and hit the ‘K’ key for “YES” if it is red or the ‘D’ key for “NO” if it is not red. The third task was One Card Learning (OCL) which requires participants to decide whether they have seen the card before or not. The ‘K’ key is hit for “YES” they had seen it before, or the ‘D’ key is hit for “NO” it had not been seen before. OCL is the longest of the three tests and in addition to processing reaction time, errors are recorded to determine accuracy of memory. These tests have also been shown sensitive to concussion (Makdissi *et al.,* 2010) and drug use. Differences in responses from different sources of impairment may ultimately provide a response signature and be useful information to develop a tool which can determine source of impairment as well as magnitude of impairment. This was not however, under investigation with this research.

The second area that Dawson and Reid's work became affiliated with is the quantification scale of fatigue-impairment. It has been built into an elaborate model (Hursh *et al.,* 2004) and used as a valid predictor of mental effectiveness, or impairment, based on daily sleep. To monitor sleep/wake cycles and feed the model with accurate data, actigraphs are used (Hursh *et al.*, 2008). Actigraphs, similar in form to a wrist watch, consist of electronics and software and determine whether the wearer is resting/sleeping or awake based on their movements. Balkin and others worked on developing actigraph-based devices to assist managing sleep in military applications over a decade ago (Balkin *et al.,* 2003) while a concurrent effort was being put into modelling human effectiveness based on sleep. Modern actigraphs are light-weight and water-proof for continuous wear even while bathing or swimming and they carry significant battery life and memory to collect data for long periods. Figure 2 shows one of the actigraphs worn during this research.

<Insert Figure 2 here>

Unlike electroencephalographic measurements of brain states, an actigraph is non-invasive and otherwise does not require wearers to change from their regular schedules or surroundings. Data is collected as close to actual circumstances as possible, but it is also accurate to results obtained from full electroencephalographic data (Morgenthaler *et al.,* 2007). Additionally, data can be collected over multiple sleep/wake cycles to more accurately represent fatigue cycles based on any inadequate rest as well as reflect shorter term performance relative to natural oscillations of the circadian rhythm. To obtain accurate data, the units must be worn continuously through a full wake/sleep cycle of a day or multiple days. This combined with their cost and the need for sleep history for real-time assessments, detracts from their suitability for workplace use for anything but intervention assessments and study. They also require acceptance by workers to wear them, as they must be worn 24 hours per day, including monitoring outside of the work place.

Modelling Sleep and Fatigue to Predict Mental Effectiveness

Through the 1980s and 1990s the Walter Reed Army Institute of Research (WRAIR) funded development of a bio-mathematical model to provide guidance to countermeasures as part of a fatigue management initiative with military personnel (Eddy and Hursh, 2001). A conclusion reached from a 2002 global workshop was that sleep-based models were capable of predicting human performance and the best performing model was one put forward by Hursh (Van Dongen, 2004). Hursh went on to expand and develop the model and ready it for more practical real world applications. The Hursh model was initially accepted by the US Department of Defence as their warfighter fatigue model (Hursh *et al.,* 2004). Hursh's model, introduced as the Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE©) Model, was used as the core for the Fatigue Avoidance Scheduling Tool (FAST©). A patent was granted for the Hursh model in 2008 (Hursh *et al.*, 2008) and it was the core model used with this research fed by actigraph measurements.

RESEARCH METHODOLOGY

A surprising lack of knowledge associated with sleep, fatigue-impairment and its performance impact at the personal or organisational level supported the need for techniques to identify and measure fatigue-impairment levels in real time. The intent of this research was to determine if a real-time surrogate measurement for fatigue-related impairment could be developed from the cognitive tests that would be as accurate as the Hursh model's prediction of impairment. For this to happen, cognitive test results would have to have significant correlation to mental effectiveness levels determined by the actigraph-fed Hursh model. By comparing the result of a cognitive test to the individual's predicted mental effectiveness level at the same point in time from the sleep model, allows assessment of cognitive tests as a basis for real-time predictions of fatigue-related impairment. The actigraph data was not visible in real time but it generated continuous profiles that allowed data matching to the minute.

Participants and Tests

Normal sleep patterns from 100 workers was to be established by collecting data for a one-month period per participant with an actigraph to monitor the quantity and quality of sleep as well as predict mental effectiveness by the minute. Due to the demanding nature of monitoring individuals' every move for a month, volunteers were sought to participate. No one was excluded from participating and participants were taken on a first-come-first-serve basis from a notice of the research. Participants expected to receive a detailed report with analysis of their month's sleep performance and data as the sole remuneration for participating. Due to the cost of the actigraphs, funding only allowed 25 to be put into circulation and be utilized for 4 rotations over 4 months. To ensure they were fully utilized in each rotation, when needed, participants external to the construction industry were invited to participate to ensure full data collection for each rotation. The collected data from each participant would allow a look-up of predicted mental effectiveness to compare with results from cognitive tests that were to be randomly taken while wearing the actigraph per Figure 3.

<Insert Figure 3 here>

Results were graphed to show the level of sinusoidal rise and fall of mental effectiveness tied to the circadian rhythm and other modelled factors. Results came from workers on different shifts. An expected important element explaining variation between groups' performance was hours awake (Lim and Dinges, 2010). This was also extracted data from the actigraph measurements as was hours slept the night before the test. An estimate of caffeine in the system at time of testing was also tracked.

The cognitive tests used were the Detection (DET), Identification (IDN) and One Card Learning (OCL) tests provided by Cogstate (Cogstate, 2012). There was no stipulated time for conducting the tests. As the actigraphs were capable of producing readings and interpreting mental effectiveness minute-by-minute, flexibility on the timing of test taking was offered to suit individual schedules. Participants had one month to contribute test results whilst wearing the actigraph, with a target of doing at least one cognitive test per day. When participants finished the tests they would transfer the data to a database on-line, along with a record of the number, type, size and time of consumption of any caffeinated drinks taken prior to the test as caffeine was previously found to have a significant effect on alertness. Participants were also given a blank form for the month they were wearing the actigraph to keep a log of their sleep. The log was to note sleep times and naps to verify actigraph results from participants, especially those who were working abnormal shifts and sleeping abnormal times.

The actigraph units were activated via a wireless connection to a computer and thereafter immediately fit to participants. The units collected data continuously and for this research the units were to be worn for at least 30 days (a month). Fatigue Science (Fatigue Science, 2012) supplied the units and software to produce full reports based on analysis of collected data. At the end of an individual's wearing period, the actigraphs were collected and the data off-loaded wirelessly to a computer with companion programmes for full analysis of the month's sleep. Having the sleep model fed by actigraph-measurements strengthens the validity of all sleep values, especially over extended time-frames such as this study, which spanned holidays and multiple work/rest schedules. Once the data was analysed to establish the activity levels of the individuals consistent with their sleep/rest and wake cycles, all statistics associated with their sleep was then derived, graphed and reported. This data fed the SAFTE© effectiveness model to give minute-by-minute estimates of individual effectiveness over the course of the time the actigraph was worn.

Results

In four separate groups, 100 participants provided data and wore an actigraph for a combined 2320 days or about 55,680 hours. The workers ranged in age from 23 to 63 and had an average age of 40.2 years. 46 females and 54 males participated; 68 were shift workers. Of the 100 participants, 83 were from the construction sector and 62 held field-based jobs. There were 5 participants who did not have valid data from a full month wearing the actigraph due to removing it for more than 4 hours, causing a session termination. Most participants kept logs of their sleep; 14 did not. Wake periods and identified short sleeps were missed on logs.

In total, by-the-minute, 4.3 million records of results and associated data was derived. Cognitive test data was extracted from the on-line database and prior to use and analysis underwent integrity and completeness checks utilizing CogState built-in integrity checks. To ensure reasonable efforts were made and that the data fit expected norms, integrity checks included that DET accuracy be 90% or greater; IDN accuracy be 80% or greater; and OCL accuracy be 53% or greater. A total of 2321 cognitive test sessions were completed. All participants completed some cognitive tests which ranged from 5 to 81 completed tests in a month by participants. The average was 23, or about 1 per work day for the four months of testing. A session was intended to include a battery of three tests DET, IDN, and OCL but of the 2321 sessions started only 2151 were accepted and it was typically the third and longest test, OCL that was abandoned or failed integrity. Associated with each participant's data were comments they provided about their test session, as well as details on caffeinated drinks consumed including timing prior to test, size and type of beverage. Each result from a cognitive test was time stamped to the minute which allowed a look-up in the 4.3 million lines of the model-determined effectiveness level at the same time per Table 1.

<Insert Table 1 here>

The 68 shift workers averaged 6.9 ± 0.95 hours of sleep per night and non-shift workers averaged 7.07 ± 0.86 hours sleep per night. Table 2 summarizes statistics for the group based on gender.

<Insert Table 2 here>

The 54 males were slightly older than the females averaging 43 years ± 10 years versus 37 years ± 11 years. The males averaged 6.6 ± 0.9 hours sleep per night and females averaged 7.4 ± 0.8 hours of sleep per night over the full study. The average mental effectiveness of the workers was 82 percent ± 12. Mental effectiveness ranged from 35.8 percent to 100 percent at time of testing. Females recorded higher mental effectiveness levels during testing than males at 85 percent ± 11 versus 78 percent ± 12. Of the 2151 valid tests there were 21 that recorded with a companion effectiveness level of 50 percent or less.

Caffeine was a commonly used stimulant amongst the workers participating. Of 100 participants, only 19 did not record any consumption of caffeinated drinks near their testing time. However, the remainder recorded a total consumption of 120,815 mg of caffeine during the period of this study, an average of 1492 mg per drinker. Coffee was the beverage of choice contributing over 95% of the caffeine consumed and estimated to be in the system at time of test followed by tea at 3%, colas at 1.9% and chocolate at 0.1%. At the time of testing, participants averaged 58 mg of caffeine in their system with males' content slightly more than females' at 61 versus 54 mg.

The hours slept and hours awake were monitored by the minute and taken from the time-stamped actigraph results. At the time of testing, participants' average time awake was 504 minutes ± 332 minutes. Females were awake slightly less time at 487 minutes ± 321 minutes versus 518 minutes ± 340 minutes for males. The females also collected more sleep before testing at 436 minutes ± 98 minutes versus 372 minutes ± 102 minutes for the males.

Of the three cognitive tests used, the DET task had the lowest mean response time and was the least variant at 309 ± 87 ms. Males displayed slightly faster response times with the DET task than females, at 318 ± 87 versus 323 ± 95 ms. The OCL task displayed the highest mean response time and was the most varied at 861 ± 223 ms. The IDN task averaged 486 ± 87 ms. These appear consistent with expectations tied to simple, choice and working memory reaction times. The IDN task was the only one in which males averaged a slower time than females, but were less varied at 505 ± 87 versus 490 ± 111 ms.

The OCL test which took about 4 minutes to complete compared to about a minute and a half for each of the other two cognitive tests had a higher percentage of incomplete tests. 2.5 percent of the OCL attempts were not completed versus 1 percent for both IDN and DET. Each of the variables monitored during the research was plotted against average alertness levels (mental effectiveness), with a trend line for each participant to assist identification of any notable patterns.

Analysis and Observations

There were no observed trends tied to individual work roles or work activity. Caffeine showed a slight upward trend relative to alertness meaning it increased with more caffeine in the system. The plot of participant age against alertness displayed a broad ‘u-shape’. It was higher amongst the 20 – 30 year age group; dropped for the 30 – 50 year age group then climbed thereafter. There is no obvious explanation for this. Time awake before test showed a negative relationship with alertness levels. This was expected; the longer one is awake, the larger is the drop in average alertness or mental effectiveness levels. Sleep before the test showed a strong positive relationship with alertness levels. This was expected; less sleep reduces alertness and mental effectiveness levels.

The DET and IDN reaction time plotted against average mental effectiveness levels did not exhibit any strong trend and the OCL processing times plotted against mental effectiveness had a broad ‘u-shape’ similar to participant age. All results exhibited a small number of outlying results with apparent data variability.

Linear Mixed Effects Models

The data from the cognitive testing and corresponding mental effectiveness derived from the actigraph data, was assembled into a single file for analysis of correlation and fit and further predictive modelling. The purpose of the analysis was to understand whether any significant relationship existed between the results of the cognitive tests and the estimated mental effectiveness results from the SAFTE© model. To assess the effects from a group of 100 with repeated measurements and with each contributing different numbers of measurements as well as assess the other variables' influence, linear mixed effects models (Faraway, 2006) was used. Fixed effects are the unknown constants that are estimated parameters for the effect of covariants such as the test results, sleep, and caffeine. The random effect is a variable with estimated parameters that describes the distribution of the random effect (Equation 1).

$$\mathbf{y_{ij} \sim X\beta + Zb + \varepsilon_{ij}} \quad \textbf{Equation 1}$$

In this case, the dependent variable ‘y’ (estimated effectiveness results) is shown to be a linear function of the multivariate measurements and estimated parameters where y_{ij} are the estimated effectiveness results estimated for individual *i* at time *j*. X is the matrix of variables that represents the systemic or fixed effects, and β is its coefficient vector. Z is a matrix of random effects variables with b the random effects. 'ε_{ij}' is the error component.

Therefore the generalized model is:

$y_{ij} = \alpha + \beta_i + \zeta_{ij} + \varepsilon_{ij}$ Equation 2

Making up the data, were 2151 results from 100 participants. Thus, $1 \le i \le 100$ and $1 \le j \le n_i$ with $\sum_{1 \le i \le 100} n_i = 2151$, y_{ij} represents the j^{th} observation of effectiveness from the i^{th} subject. The term ζ_{ij} represents the fixed effects terms including the effects of caffeine, time awake before testing, sleep obtained before testing, age and gender.

The random effects term is assumed to follow a normal distribution with a mean that reflects the overall aggregated baseline response that allows the dependence between measurements from the same participants to be acknowledged. The variance on the random term reflects the heterogeneity in responses between subjects and allows separate intercept terms to be fitted for each participant. Any factors not measured which may have contributed to individual effectiveness levels will be contained within the random term, as will any measurement error associated with the measured values.

The analysis was implemented using the lme4 package within the R statistical package (R Development Core Team, 2010). Several linear mixed effects (LME) models were derived to help analyse the relationship between the cognitive test results and the mental effectiveness output from the fatigue-model with actigraph measurements as input, whilst accounting for normal workplace changes in alertness (caffeine and other factors). Table 3 illustrates the predictive models that were initially analysed, comprised of variations of the variables for which data was collected during the research. Entirely fixed effects models were also looked at to compare to the mixed effects results and to verify that the linear mixed effects approach was proper and valid.

<Insert Table 3 here>

Each model had a common response variable ‘Effectiveness’ of the subject at the point of taking the cognitive test as derived from the actigraph-fed fatigue model. Common independent variables of interest, including the subject, caffeine content when testing, time awake before testing, amount of sleep prior to testing, age and gender, were included in the model to help explain the measured ‘effectiveness’ results. Determining whether there was any significant relationship between the results from the cognitive tests for DET, IDN and OCL and the sleep-based fatigue model result was of primary interest. Both the processing speeds as well as the accuracy measurements from the tests were considered when analysing results.

The three cognitive test results were analysed to understand the influence of all factors on the effectiveness results as shown in Table 4. The effect of the cognitive test on the outcomes of 3 tests, DET, OCL, and IDN is shown.

<Insert Table 4 here>

There are five levels displayed for each outcome that build on a baseline model showing (i) subject effect which has a random intercept term for the participant; (ii) the addition of the effect of a test result, and; (iii) the additional effect from caffeine, hours slept prior to test, hours awake before test, participant age and participant gender; (iv) the combined effect of (ii) and (iii); (v) the effect of adding the results from the other two tests. In all cases, the results are considered significant at the 5% significance level which from the table is illustrated by an (*). All but one result in Table 4 far exceeded this and are highly significant at the 2.2e-16 level as shown in the Pr(>Chisq) column.

The results shown illustrate that combining the results of any of the three cognitive tests makes a statistically significant stronger model for predicting mental effectiveness. To predict the 'effectiveness' level with cognitive test results, OCL would be the strongest of the three tests (strongest correlation) followed by DET then IDN based on the 'Chisq' values shown (Model D followed by C followed by B). Additionally the influence of caffeine, hours of sleep before the test, hours awake before the test, age and gender are shown. Hours of sleep before the test showed most effect for all tasks meaning that the less sleep one had before the test, the worse was their result, classified as 'effectiveness'. This was followed by time awake before the test was taken. Caffeine had the least effect of the three.

Accuracy of Results

Table 5 shows the ANOVA results with the influence of accuracy from each test included. <Insert Table 5 here>

Each cognitive test maintained a count of the number of errors made, as well as the processing time which was the measurement used for processing accuracy. As a factor, accuracy was a much weaker predictor of 'effectiveness' than processing time for each test. However, it was a significant factor in each case; weakest for the IDN test and strongest for OCL test per the 'Pr(>Chisq)' values of Table 5.

The statistical significance relative to the results shows that the cognitive screening tests strongly match with changes in predicted individual mental effectiveness determined by the SAFTE© model from measured sleep/wake cycles. The cognitive tests displayed a strong correlation to the 'mental effectiveness' levels in the presence of confounding elements, some of which were identified and built into the assessment and some which were not and make up the error profile. This suggests that multiple terms could be used and combined to create a cognitive test fatigue-impairment predictive model that would closely match the results of predicted 'mental effectiveness' from a fatigue model fed by actigraph measurements. There are mathematical tools to assist in determining which terms are better than others for a predictive model such as the Akaike Information Criterion ($AIC = -2\ln(f) + 2p$) which is a technique that seeks a balance to adding parameters for maximum log likelihood by considering the strength of what they add. What is sought is a minimum AIC value. To develop a usable fatigue-impairment model for the workplace utilizing the cognitive tests, simplicity versus accuracy would be a consideration therefore additional parameters must greatly improve the likelihood of the model to reduce the AIC values.

Combining different parameters into a model in this fashion and checking their influence suggests a good performing cognitive-test based fatigue-impairment model would be derived from the parameters OCL, OCL_err, time awake and amount of sleep. A possible fatigue-impairment model based on a cognitive test is shown as Equation 3.

Effectiveness ~ Subject + OCL + OCL_err + Sleep + Awake **Equation 3**

The results of this research validated a new exciting opportunity for real-time screening for fatigue-related impairment in the workplace. It also validated that cognitive tests have no significant operational hurdles for their use. Additionally, factors other researchers (Falleti *et al.,* 2003, Collie *et al.,* 2007, Maruff *et al.,* 2005, Lim and Dinges, 2010)have found as strong influencers on alertness, such as stimulants and hours awake, were addressed and built into these results with measured influence. This study took workers as they presented themselves in real workplaces with all individual variability they normally brought into the workplace and gave results correlating to mental effectiveness over the course of their shifts. The results give confidence cognitive tests could be used to develop predictive models of mental effectiveness as surrogate measurements of fatigue-impairment. With the knowledge that these cognitive tests highly correlate with fluctuating mental effectiveness levels, predictive models could be further fine-tuned to individual workers by collecting their specific response profiles resulting in tighter estimates of their fatigue-impairment expressed as a Blood Alcohol Concentration level. Estimating fatigue-impairment for workers who either appear impaired, have worked long hours in a day or week or are working an unusual shift with very early start times or very late work hours would be practical target candidates for tests.

CONCLUSIONS

While the results from the simple real-time tests correlated strongly with a comprehensive and validated actigraph-fed fatigue model, it must be remembered that both approaches carry error and are just estimates not the penultimate measurement of an individual’s state of mind.

Two of the variables found to have significant relationships with the test results and which were part of the strongest predictive fatigue-impairment model with cognitive test results, were the amount of sleep individuals had before testing and the time they were awake before testing. Both of these measurements were derived from actigraph measurements as opposed to an individual estimating these values. A model utilizing these factors to estimate an individual's 'mental effectiveness' with cognitive tests but without actigraphs would be less accurate and have additional error. As sleep in particular, accounting for sleep quality, is poorly estimated, this factor would need to be reconsidered in a working cognitive test model. However, a valid fatigue-impairment model not containing these variables and just relying on cognitive test results could also be selected.

To take full advantage of the findings of this research, the development and implementation of any resulting cognitive test predictive model can only be considered a part of a working solution. It provides one missing piece of a strategy for addressing fatigue in the industry. It will require commitment to several other changes including implementing a fatigue risk management system but it gives invaluable protection by monitoring for fatigue impairment.

References

Balkin, T. J., Belenky, G. L., Hall, S. W., Kamimori, G. H., Redmond, D. P., Sing, H. C., Thomas, M. L., Thorne, D. R. and Wesensten, N. J. (2003), "System and method for predicting human cognitive performance using data from an actigraph", U.S. Patent No. 6527715 B2,Washington, DC: U.S. Patent and Trademark Office.

Beaulieu, J. K. (2005), "The issues of fatigue and working time in the road transport sector", working paper 92-2-117720-3,International Labour Office, Geneva, 2005.

Cogstate, "Home page", available at: http://cogstate.com/, (accessed 12 August 2014)

Collie, A., Maruff, P., McStephen, M. and Darby, D. G. (2003), "Psychometric issues associated with computerised neuropsychological assessment of concussed athletes", *Br J Sports Med*, 37(6), 556-9.

Collie, A., Makdissi, M., Maruff, P., Bennell, K. and McCrory, P. (2006), "Cognition in the days following concussion: Comparison of symptomatic versus asymptomatic athletes", *J Neurol Neurosurg Psychiatry*, 77(2), 241-5.

Collie, A., Maruff, P., Makdissi, M., McCrory, P., McStephen, M. and Darby, D. G. (2003), "Cogsport: Reliability and correlation with conventional cognitive tests used in post-concussion medical examinations", *Clinical Journal of Sport Medicine*, 13, 28-32.

Collie, A., Darekar, A., Weissgerber, G., Toh, M. K., Snyder, P. J., Maruff, P. and Huggins, J. P. (2007), "Cognitive testing in early-phase clinical trials: Development of a rapid computerized test battery and application in a simulated phase 1 study", *Contemp Clin Trials*, 28(4), 391-400.

Darby, D., Maruff, P., Collie, A. and McStephen, M. (2002), "Mild cognitive impairment can be detected by multiple assessments in a single day", *Neurology*, 59(7), 1042-6.

Darby, D. G., Brodtmann, A., Pietrzak, R. H., Fredrickson, J., Woodward, M., Villemagne, V. L., Fredrickson, A., Maruff, P. and Rowe, C. (2011), "Episodic memory decline predicts cortical amyloid status in community-dwelling older adults", *J Alzheimers Dis*, 27(3), 627-37.

Dawson, D. and Reid, K. (1997), "Fatigue, alcohol and performance impairment", *Nature*, Vol 338, 235-7.

Dinges, D. F., Graeber, R. C., Rosekind, M. R., Samel, A. and Wegmann, H. M. (1996), "Principles and guidelines for duty and rest scheduling in commercial aviation", in, *Nasa*, (Ed.), Moffett Field, California: National Aeronautics and Space Administration, 15.

Dinges, D. F., Pack, F., Williams, K., Gillen, K. A., Powell, J. W., Ott, G. E., Aptowicz, C. and Pack, A. I. (1997), "Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night", *Sleep*, 4(Apr 20), 267-77.

Eddy, D.R. and Hursh, S.R. (2001). "Fatigue Avoidance Scheduling Tool: Phase I SBIR Final report", F41624-99-C-6041, Brooks Air Force Base, TX: Air Force research Laboratory, Human Effectiveness Directorate, Biodynamic and Protection Division, Flight Motion Effects Branch.

Ellis, K. A., Bush, A. I., Darby, D., De Fazio, D., Foster, J., Hudson, P., Lautenschlager, N. T., Lenzo, N., Martins, R. N., Maruff, P., Masters, C., Milner, A., Pike, K., Rowe, C., Savage, G., Szoeke, C., Taddei, K., Villemagne, V., Woodward, M. and Ames, D. (2009), "The australian imaging, biomarkers and lifestyle (aibl) study of aging: Methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of alzheimer's disease", *Int Psychogeriatr*, 21(4), 672-87.

Falleti, M. G., Maruff, P., Collie, A. and Darby, D. G. (2006), "Practice effects associated with the repeated assessment of cognitive function using the cogstate battery at 10-minute, one week and one month test-retest intervals", *J Clin Exp Neuropsychol*, 28(7), 1095-112.

Falleti, M. G., Maruff, P., Collie, A., Darby, D. G. and McStephen, M. (2003), "Qualitative similarities in cognitive impairment associated with 24 h of sustained wakefulness and a blood alcohol concentration of 0.05%", *J Sleep Res*, 12(4), 265-74.

Faraway, J. J. (2006), *Extending the linear model with r: Generalized linear, mixed effects and nonparametric regression models*. Taylor & Francis Group, Boca Raton.

Fatigue Science, *Home page*, available at: http://www.fatiguescience.com/home. (accessed 22 September 2014).

Fletcher, A., Lamond, N., van den Heuval, C. J. and Dawson, D. (2003), "Prediction of performance during sleep deprivation and alcohol intoxication using a quantitative model of work-related fatigue", *Sleep Research Online*, 5(2), 67-75.

Fredrickson, J., Maruff, P., Woodward, M., Moore, L., Fredrickson, A., Sach, J. and Darby, D. (2010), "Evaluation of the usability of a brief computerized cognitive screening test in older people for epidemiological studies", *Neuroepidemiology*, 34, 65-75.

Hursh, S. R., Elsmore, T. and Eddy, D. R. (2008), "Interface for a system and method for evaluating task effectiveness based on sleep pattern", U.S.Patent No. 7,384,394, Washington, DC: U.S. Patent and Trademark Office.

Hursh, S. R., Redmond, D. P., Johnson, M. L., Thorne, D. R., Belenky, G., Balkin, T. J., Storm, W. F., Miller, J. C. and Eddy, D. R. (2004), "Fatigue models for applied research in warfighting", *Aviation, Space, and Environmental Medicine*, 00(0), 1-10.

Lamond, N. and Dawson, D. (1999), "Quantifying the performance impairment associated with fatigue", *Journal of Sleep Research.*, 8(4), 255-62.

Levine, B., Roehrs, T., Zorick, F. and Roth, T. (1988), "Daytime sleepiness in young adults", *Sleep*, 11(1), 39-46.

Lim, J. and Dinges, D. F. (2010), "A meta-analysis of the impact of short-term sleep deprivation on cognitive variables", *Psychological Bulletin*, 136(3)(May 2010), 375-89.

Makdissi, M., Darby, D., Maruff, P., Ugoni, A., Brukner, P. and McCrory, P. R. (2010), "Natural history of concussion in sport; markers of severity and implications for management", *Am J Sports Med,*, 38(3), 464-71.

Makdissi, M., Collie, A., Maruff, P., Darby, D. G., Bush, A., McCrory, P. and Bennell, K. (2001), "Computerised cognitive assessment of concussed australian rules footballers", *Br J Sports Med*, 35(5), 354-60.

Maruff, P., Collie, A., Darby, D., Weaver-Cargin, J. and McStephen, M. (2004), "Subtle memory decline over 12 months in mild cognitive impairment", *Dement Geriatr Cogn Disord*, 18(3), 342-8.

Maruff, P., Falleti, M. G., Collie, A., Darby, D. G. and McStephen, M. (2005), "Fatigue-related impairment in speed, accuracy and variability of psychomotor performance: Comparison with blood alcohol levels", *Journal of Sleep Research*, 14, 21-7.

Maruff, P., Thomas, E., Cysique, L., Brew, B., Collie, A., Snyder, P. and Pietrzak, R. H. (2009), "Validity of the cogstate brief battery: Relationship to standardized tests and sensitivity to cognitive impairment in mild traumatic brain injury, schizophrenia, and aids dementia complex", *Arch Clin Neuropsychol*, 24(2), 165-78.

Morgenthaler, T., Alessi, C., Friedman, L., Owens, J., Kapur, V., Boehlecke, B., Brown, T., Chesson, A., Coleman, J., Lee-Chong, T., Pancer, J. and Swick, T. (2007), "Practice parameters for the use of actigraphy in the assessment of sleep and sleep disorders: An update for 2007", *Sleep*, Vol. 30(No. 4).

Moriarity, J., Collie, A., Olson, D., Buchanan, J., Leary, P., McStephen, M. and McCrory, P. (2004), "A prospective controlled study of cognitive function during an amateur boxing tournament", *Neurology*, 62(9), 1497-502.

National Institute for Occupational Safety and Health, "Construction safety", available at: http://www.cdc.gov/niosh/topics/constructionsafety/. (Accessed 12 September 2012).

Pietrzak, R. H., Olver, J., Norman, T., Piskulic, D., Maruff, P. and Snyder, P. J. (2009), "A comparison of the cogstate schizophrenia battery and the measurement and treatment research to improve cognition in schizophrenia (matrics) battery in assessing cognitive impairment in chronic schizophrenia", *J Clin Exp Neuropsychol*, 1-12.

Pietrzak, R. H., Snyder, P. J., Jackson, C. E., Olver, J., Norman, T., Piskulic, D. and Maruff, P. (2009), "Stability of cognitive impairment in chronic schizophrenia over brief and intermediate re-test intervals", *Hum Psychopharmacol*, 24(2), 113-21.

Powell, R. I. (2009), "Fatigue-impairment in British Columbia's construction industry", *In: Proceedings of the 27th International System Safety Conference-The Win Win Solution*, 3 - 7 August, 2009, International System Safety Society, Huntsville, Alabama, Vol. August 3 - 7, 2009.

Powell, R. and Copping, A. G. A. (2010), "Sleep deprivation and its consequences in construction workers", *Journal of Construction Engineering and Management*, 136(10), 1086-92.

R Development Core Team (2010), "R: A language and environment for statistical computing", R Foundation for Statistical Computing, Vienna, Austria.

Reason, J. (2008), *The human contribution : Unsafe acts, accidents and heroic recoveries*, Ashgate Publishing Limited, Surrey.

Rosekind, M. R., Co, E. L., Gregory, K. B. and Miller, D. L. (2000), "Crew factors in flight operations XIII: A survey of fatigue factors in corporate/executive aviation operations", Technical Memorandum #209610, National Technical Information Service, NASA Moffett Field, CA, 2000.

Roth, T. "Dealing with excessive sleepiness: Therapeutics and public health considerations", available at: http://www.medscape.com/viewprogram/4394 (accessed 11 November 2012)

Snyder, P., Bednar, M., Cromer, J. and Maruff, P. (2005), "Reversal of scopolamine-induced deficits with a single dose of donepezil, an acetylcholinesterase inhibitor", *Alzheimer's & Dementia*, 1, 126-35.

Snyder, P., Werth, J., Giordani, B., Caveney, A. F., Feltner, D. and Maruff, P. (2005), "A method for determining the magnitude of change across different cognitive functions in clinical trials: The effects of acute administration of two different doses alprazolam", *Hum Psychopharmacol*, 20(4), 263-73.

Straume-Naesheim, T. M., Andersen, T. E., IM, K. H., McIntosh, A. S., Dvorak, J. and Bahr, R. (2009), "Do minor head impacts in soccer cause concussive injury? A prospective case-control study", *Neurosurgery*, 64(4), 719-25; discussion 25.

Strauss, S. (2003), "Pilot fatigue", *available at: http://www.europeanpilotacademy.com/wp-content/uploads/2013/11/Pilot-Fatigue.pdf (accessed 12 August 2014)*

Van Dongen, H., Rogers, N. L. and Dinges, D. F. (2003), "Sleep debt: Theoretical and empirical issues", *Sleep and Biological Rhythms*, 1(1), 5-13.

Van Dongen, H. P. A. (2004), "Comparison of mathematical model predictions to experimental data of fatigue and performance", *Aviation, Space and Environmental Medicine*, 75(No. 3), A15 - 36.

Wehr, T. A., Moul, D. E., Barbato, G., Giesen, H. A., Seidel, J. A., Barker, C. and Bender, C. (1993), "Conservation of photoperiod-responsive mechanisms in humans", *American Journal of Physiology*, Vol 265(Issue 4), 846-R57.

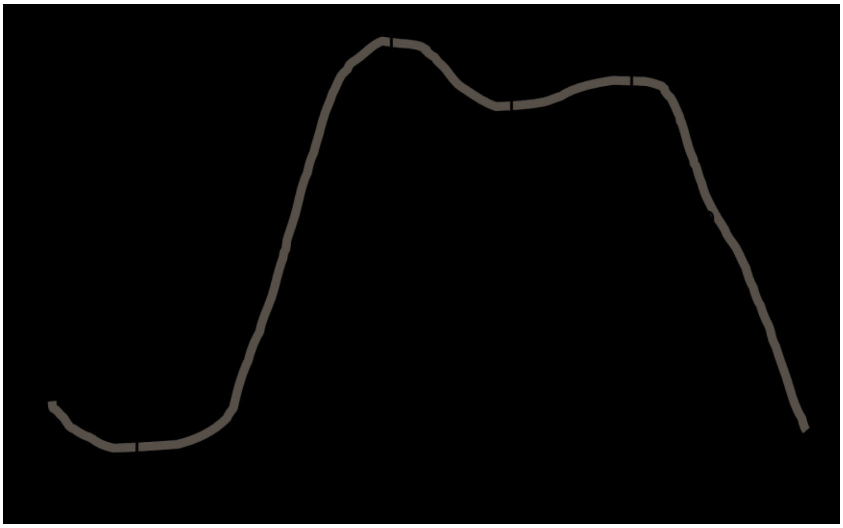

Figure 1 - Profile of alertness governed by the circadian cycle.

Figure 2 - Wrist-worn actigraph from Fatigue Science

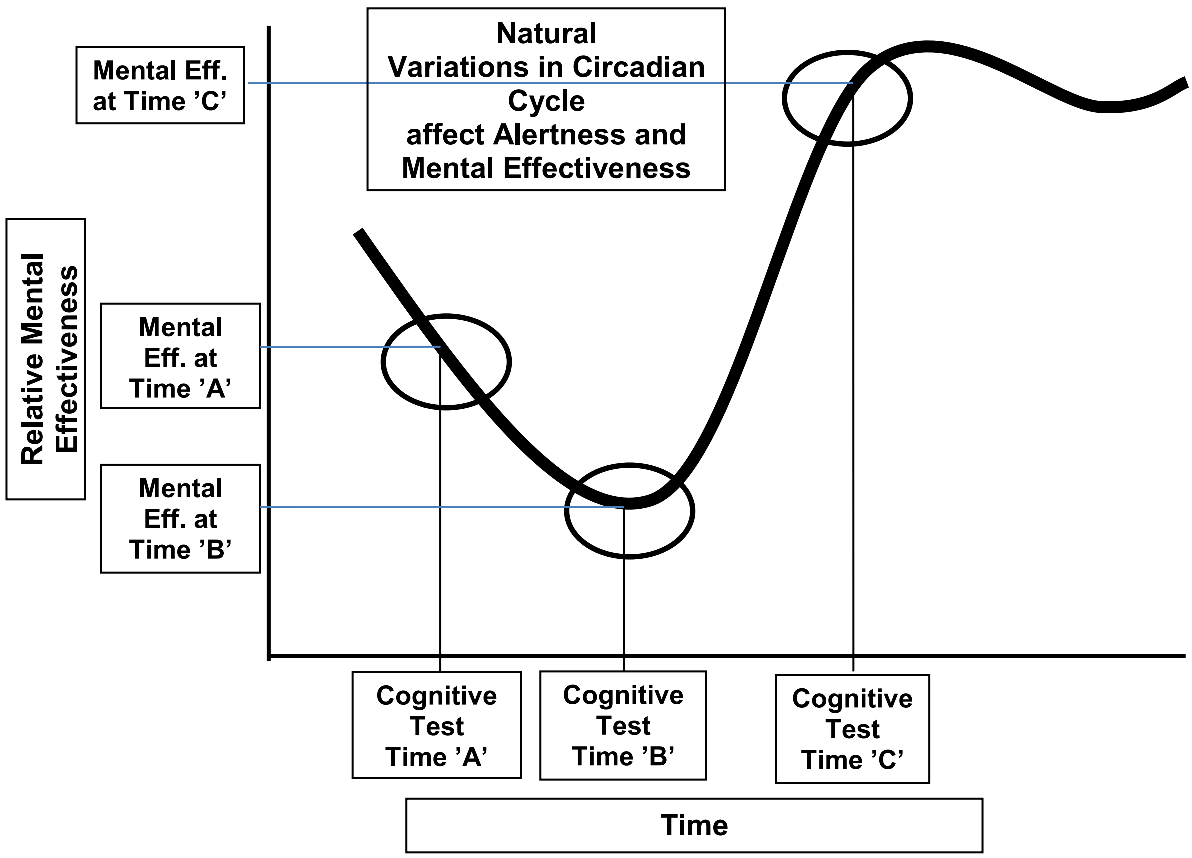

Figure 3 – Expected Circadian-based variations in Mental Effectiveness

Table 1 - Example of data derived from each participant's effectiveness profile.

Zulu Time	Local Time	Actigraph	In Bed (Up=0 Bed=1 Bad=2)	Sleep Wake (Sleep=0 Awake=1)	Mental Effectiveness
13/10/2011 20:47	13/10/2011 13:47	347	0	1	82.63
13/10/2011 20:48	13/10/2011 13:48	212	0	1	82.55
13/10/2011 20:49	13/10/2011 13:49	117	0	1	82.46
13/10/2011 20:50	13/10/2011 13:50	198	0	1	82.38
13/10/2011 20:51	13/10/2011 13:51	63	0	1	82.29
13/10/2011 20:52	13/10/2011 13:52	195	0	1	82.21
13/10/2011 20:53	13/10/2011 13:53	203	0	1	82.12
13/10/2011 20:54	13/10/2011 13:54	200	0	1	82.04

Table 2 - Summary Statistics by Gender

Variable	54 Males			46 Females			ALL 100		
	Avg	Std Dev	Range	Avg	Std Dev	Range	Avg	Std Dev	Range
Age (years)	43	10	24 – 60	37	11	23 - 63	40	11	23 - 63
Caffeine (mg)	61	94	0 - 800	54	91	0 - 1125	58	92	0 - 1125
Mental Effectiveness (%)	78	12	36- 100	85	11	50.7 - 100	82	12	36- 100
Avg Sleep Full Study (hours)	6.6	0.9	5.1 – 9.7	7.4	0.8	5.9 – 9.2	7.0	0.9	5.1 – 9.7
Awake Before Test (min)	518	340	5 – 2092	487	321	5 - 1716	504	332	5 - 2092
Sleep Before Test (min)	372	102	0 – 835	436	98	100 - 810	401	105	0 - 835
IDN Response	502	88	330 – 1054	490	111	333 - 1370	497	99	330 - 1370
DET Response	315	86	192 – 1834	323	95	200 - 1139	319	90	192 - 1834
OCL Response	862	186	404 -1877	863	248	382 - 2727	863	220	382 - 2727

Table 3 - Initial Models Analysed with Linear Mixed Effects modelling

Model	Response Variable	Random Effects	Fixed Effects
A	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender*
B	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender+IDN
C	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender+DET
D	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender+OCL
E	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender+ IDN+DET+OCL
F	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender+IDN accuracy
G	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender+DET accuracy
H	Effectiveness	Subject	Caffeine+Awake+SBT+Age+Gender+OCL accuracy

*Awake = Time awake before taking the cognitive test
*SBT = Sleep before taking the cognitive test

ANOVA Results					
Test	**Model**	**Log-Likelihood**	**Chisq**	**df**	**Pr(>Chisq)**
DET	Subject level random effect	-7846.3			
	Additional effect of DET	-7765.6	161.39	1	<2.2e-16***
	Additional Effect of Caffeine + Sleep + Hours awake + Age+ Gender (no DET)	-7209.7	1111.73	4	<2.2e-16***
	Additional Effect of Caffeine + Sleep + Hours awake + Age+ Gender + DET	-7138.4	142.71	1	<2.2e-16***
	Additional effect + OCL+IDN	-5111.4	4053.96	2	<2.2e-16***
OCL	Subject level random effect	-7846.3			
	Additional effect of OCL	-5574.8	-5543	1	<2.2e-16***
	Additional Effect of Caffeine + Sleep + Hours awake+Age+Gender (no OCL)	-7209.7	0	4	1
	Additional Effect of Caffeine + Sleep + Hours awake + Age + Gender + OCL	-5193.2	4033.05	1	<2.2e-16***
	Additional effect + DET+IDN	-5111.4	163.63	2	<2.2e-16***
IDN	Subject level random effect	-7846.3			
	Effect of IDN	-7774.6	143.46	1	<2.2e-16***
	Additional Effect of Caffeine + Sleep + Hours awake + Age + Gender	-7209.7	1129.7	4	<2.2e-16***
	Additional Effect of Caffeine + Sleep + Hours awake + Age + Gender + IDN	-7148.8	121.9	1	<2.2e-16***
	Additional effect + DET+OCL	-5111.4	4074.8	2	<2.2e-16***
Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Table 4 - Correlation of Cognitive test on estimated Effectiveness

Test	Model	Log-Likelihood	Chisq	df	Pr(>Chisq)
DET	Base = Effect of Caffeine + Sleep + Hours awake + Age+ Gender	-7209.7			
	Plus effect of DET	-7138.4	142.7	1	< 2.2e-16 ***
	Plus effect of DET_err	-7134.3	8.2	1	0.004274**
OCL	Base = Effect of Caffeine + Sleep + Hours awake + Age + Gender	-7209.7			
	Plus effect of OCL	-5193.2	4033	1	< 2.2e-16 ***
	Plus effect of OCL_err	-5177.1	32.3	1	<1.33e-08***
IDN	Base = Effect of Caffeine + Sleep + Hours awake + Age + Gender	-7209.7			
	Plus effect of IDN	-7148.8	121.9	1	< 2e-16 ***
	Effect of IDN_err	-7147.2	3.2	1	0.0724*
Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Table 5 - ANOVA Results from Accuracy Measurements